

## Постановка задачі

У результаті проведення експерименту по вивченню 2-х властивостей було обстежено N об'єктів. Результати вимірів властивостей представлені в таблиці. Необхідно провести

- А) ізотропне;
- Б) ізотонічне;
- В) ізоморфічне

перетворення отриманої сукупності даних. По методу куль розбити досліджувану сукупність об'єктів на підмножини.

Як результат представити:

- А) матрицю відстаней , отриману для кожного з видів перетворення;
- Б) критичне значення радіуса кулі і сукупність отриманих підмножин.

Вихідні дані приведені в табл. 1.

Таблиця 1.

X	Y
1.2	4.2
1.3	4.4
1.4	4.8
1.5	5.1
1.7	6.8
2.2	6.6
2.3	7.8
2.6	4.9
2.7	5.2
2.9	5.3

## Виконання завдання

### А) Ізотропне перетворення

Найчастіше використовуються такі таксономічні методи, за допомогою яких виділяються підмножини, однорідні в змісті ізотропності (об'єкти, що належать тій самій підмножині, мало відрізняються друг від друга за рівнем і структурою значень ознак). Справа в тім, що, як правило, позбутися від одиниць виміру ознак можна шляхом так званої стандартизації:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (1.1)$$

де  $x_{ij}$  –  $i$ -та реалізація  $j$ -тої ознаки;

$\bar{x}_j$  – середнє значення  $j$ -тої ознаки;

$s_j$  – середньоквадратичне відхилення  $j$ -тої ознаки.

У результаті обидві компоненти реалізації ознак (структура і рівень) вирівняні.

Стандартизовані значення:

Стандартизовані дані	
X	Y
-1.230557	-1.123821
-1.072793	-0.952245
-0.915029	-0.609094
-0.757266	-0.35173
-0.441738	1.106663
0.3470801	0.935088
0.5048438	1.964542
0.9781349	-0.523306
1.1358985	-0.265942
1.4514259	-0.180154

Відстань між об'єктами обчислюється за формулою звичайної евклідової метрики

$$d_{ij} = \sqrt{\sum_{k=1}^m (x_{ik} - x_{jk})^2} \quad (1.2)$$

Матриця відстаней наведена в табл. 2.

Таблиця 2

Матриця відстаней										
	d1	d2	d3	d4	d5	d6	d7	d8	d9	d10
d1	0.0000	0.2331	0.6037	0.9056	2.3659	2.5938	3.5425	2.2889	2.5172	2.8432
d2	0.2331	0.0000	0.3777	0.6784	2.1534	2.3618	3.3161	2.0953	2.3129	2.6397
d3	0.6037	0.3777	0.0000	0.3019	1.7798	1.9943	2.9393	1.8951	2.0794	2.4050
d4	0.9056	0.6784	0.3019	0.0000	1.4921	1.6957	2.6378	1.7439	1.8951	2.2153
d5	2.3659	2.1534	1.7798	1.4921	0.0000	0.8073	1.2775	2.1617	2.0912	2.2891
d6	2.5938	2.3618	1.9943	1.6957	0.8073	0.0000	1.0415	1.5891	1.4369	1.5695
d7	3.5425	3.3161	2.9393	2.6378	1.2775	1.0415	0.0000	2.5325	2.3180	2.3443
d8	2.2889	2.0953	1.8951	1.7439	2.1617	1.5891	2.5325	0.0000	0.3019	0.5846
d9	2.5172	2.3129	2.0794	1.8951	2.0912	1.4369	2.3180	0.3019	0.0000	0.3270
d10	2.8432	2.6397	2.4050	2.2153	2.2891	1.5695	2.3443	0.5846	0.3270	0.0000

Критичне значення радіуса кулі:

r=
1,041473

Таким чином отримаємо такі однорідні підмножини:

$A1=\{d1,d2,d3,d4\}$

$A2=\{d5,d6,d7\}$

$A3=\{d8,d9,d10\}$

Для аналізу отриманих результатів запишемо найбільше та найменше значення кожної ознаки та середнє значення ознак для кожної однорідної групи об'єктів (табл. 4).

Таблиця 4

	X	Y
min	1.2	4.2
max	2.9	7.8
середнє	1.98	5.51
сер для A1	1.35	4.63
сер для A2	2.07	7.07
сер для A3	2.73	5.13

Отже, однорідні групи характеризуються такими властивостями:

A1 Низькі витрати на одиницю продукції, низька фондовіддача.

A2 Середні витрати на одиницю продукції, висока фондовіддача.

A3 Високі витрати на одиницю продукції, низька фондовіддача

Б) Ізотонічне перетворення:

Перш ніж почати виділення ізотонічних підмножин, потрібно позбутися від одиниць виміру ознак. З цією метою використовуємо перетворення, що виключить зі значень кожної ознаки компонентів структури, але збереже компонентів рівня. Для цього перетворимо ознаки  $X_i = [x_{1i}, x_{2i}, \dots, x_{ni}]^T$

до виду

$$V_i = [v_{1i}, v_{2i}, \dots, v_{ni}]^T, \quad (1.3)$$

за допомогою ізотонічного перетворення (Таблиця 5)

$$v_{ji} = \frac{x_{ji}}{\sum_{j=1}^m x_{ji}} \quad (1.4)$$

Таблиця 5- значення кожної ознаки без компонентів структури.

V1	V2
0.060606	0.076225
0.065657	0.079855
0.070707	0.087114
0.075758	0.092559
0.085859	0.123412
0.111111	0.119782

0.116162	0.141561
0.131313	0.088929
0.136364	0.094374
0.146465	0.096189

Отримані в такий спосіб ознаки характеризуються тим, що їхні довжини (що розуміються як довжини векторів) дорівнюють одиниці, тобто для ознак  $V_j$  виконується умова:

$$\|V_i\| = \sum_{j=1}^m v_{ji} = 1 \quad (1.5)$$

Вихідні ж ознаки звичайно мають різну довжину.

При одержанні ознак  $V_j$  виключаються одиниці виміру вихідних ознак і встановлюються однакові довжини ознак, рівні 1. Таким чином, значення  $v_{ij}$  характеризують "потенціал", "масштаб", "позицію" або "ранг" даного об'єкта в розглянутій сукупності об'єктів. У свою чергу сума значень

$$w_j = \sum_{i=1}^n v_{ji} \quad (1.6)$$

характеризує внесок значень усіх властивостей даного об'єкта в значення всіх властивостей сукупності, що включає всі об'єкти. Ця величина синтетичним образом представляє "потенціал" або "позицію"  $j$ -го об'єкта в сукупності. Вектор  $W = [w_1, w_2, \dots, w_m]^T$  називається ізотонічним показником сукупності даних (таблиця 6).

Таблиця 6 – Ізотонічні показники сукупності даних.

W=V1+V2
0.136831
0.145511
0.157821
0.168317
0.209271
0.230893
0.257722
0.220242
0.230738
0.242653

Для виділення підмножин, однорідних в ізотонічному відношенні, відстань між точками об'єктами визначається за формулою "міської" метрики:

$$d_{ij} = |w_i - w_j| \quad (1.7)$$

Елементи  $d_{ij}$  приймають значення, рівні нулеві, якщо значення показників  $w_i$  і  $w_j$  ідентичні (табл. 7).

Таблиця 7 – Матриця відстаней.

Матриця відстаней										
	d1	d2	d3	d4	d5	d6	d7	d8	d9	d10
d1	0.0000	0.0087	0.0210	0.0315	0.0724	0.0941	0.1209	0.0834	0.0939	0.1058
d2	0.0087	0.0000	0.0123	0.0228	0.0638	0.0854	0.1122	0.0747	0.0852	0.0971

d3	0.0210	0.0123	0.0000	0.0105	0.0514	0.0731	0.0999	0.0624	0.0729	0.0848
d4	0.0315	0.0228	0.0105	0.0000	0.0410	0.0626	0.0894	0.0519	0.0624	0.0743
d5	0.0724	0.0638	0.0514	0.0410	0.0000	0.0216	0.0485	0.0110	0.0215	0.0334
d6	0.0941	0.0854	0.0731	0.0626	0.0216	0.0000	0.0268	0.0107	0.0002	0.0118
d7	0.1209	0.1122	0.0999	0.0894	0.0485	0.0268	0.0000	0.0375	0.0270	0.0151
d8	0.0834	0.0747	0.0624	0.0519	0.0110	0.0107	0.0375	0.0000	0.0105	0.0224
d9	0.0939	0.0852	0.0729	0.0624	0.0215	0.0002	0.0270	0.0105	0.0000	0.0119
d10	0.1058	0.0971	0.0848	0.0743	0.0334	0.0118	0.0151	0.0224	0.0119	0.0000

Узагалі можна стверджувати, що чим вище значення  $d_{ij}$ , тим більше відрізняються між собою розглянуті об'єкти за рівнем значень властивостей. Низькі значення  $d_{ij}$  свідчать про малі розходження об'єктів за рівнем значень властивостей.

Критичне значення радіуса кулі:

$r=$
0,0151

Таким чином отримаємо підмножини:

$A1=\{d1,d2,d3,d4,d5\}$

$A2=\{d6,d7\}$

$A3=\{d8,d9,d10\}$

Для аналізу отриманих результатів запишемо найбільше та найменше значення кожної ознаки та середнє значення ознак для кожної однорідної групи об'єктів (табл. 9).

Таблиця 9.

	X	Y
min	1.2	4.2
max	2.9	7.8
середнє	1.98	5.51
сер для A1	1.42	5.06
сер для A2	2.25	7.2
сер для A3	2.73	5.13

Отже, за рівнями значень показників об'єкти розподілились так :

A1 – низькі витрати на одиницю продукції, низька фондовіддача.

A2 – середні витрати на одиницю продукції, висока фондовіддача.

A3 – високі витрати на одиницю продукції, низька фондовіддача.

В) Ізоморфічне перетворення

Такі підмножини характеризуються однорідністю з точки зору структури значень. Тобто, пропорції значень відповідних ознак будуть мало відрізнятись одна від одної. Для їх одержання потрібно провести наступні перетворення

$$z_{ij} = \frac{\frac{x_{ij}}{\sum_{i=1}^m x_{ij}}}{\sum_{j=1}^n \frac{x_{ij}}{\sum_{i=1}^m x_{ij}}} = \frac{v_{ij}}{\sum_{j=1}^n v_{ij}} \quad (1.8)$$

В результаті одержимо нові точки  $R_i = (z_{i1}, z_{i2}, \dots, z_{in})$ . Вони характеризуються тим,

що  $\|R_i\| = \sum_{j=1}^n z_{ij} = 1$ . Відстань між такими точками визначається також за допомогою звичайної евклідової метрики:

$$d_{ij} = \sqrt{\sum_{k=1}^n (z_{ki} - z_{kj})^2} \quad (1.9)$$

Мінімальна відстань між об'єктами  $R_i$  та  $R_j$  спостерігається тоді, коли вихідні вектори  $P_i$  та  $P_j$  колінеарні. Максимум відстані досягається, якщо вихідні вектори ортогональні. Помітимо, що значення перетворених векторів лежать в межах від 0 до 1. Отже в такому випадку максимальна відстань буде становити  $\sqrt{2}$ .

Проведемо стандартизацію даних і запишемо їх до табл. 10, а матрицю відстаней – в табл. 11.

Таблиця 10 – Стандартизовані дані.

V1	V2	W=V1+V2	Z1=V1/W	Z2=V2/W
0.060606	0.076225	0.136831	0.442926	0.557074
0.065657	0.079855	0.145511	0.451213	0.548787
0.070707	0.087114	0.157821	0.44802	0.55198
0.075758	0.092559	0.168317	0.45009	0.54991

0.085859	0.123412	0.209271	0.410276	0.589724
0.111111	0.119782	0.230893	0.481223	0.518777
0.116162	0.141561	0.257722	0.450724	0.549276
0.131313	0.088929	0.220242	0.596221	0.403779
0.136364	0.094374	0.230738	0.59099	0.40901
0.146465	0.096189	0.242653	0.603596	0.396404

Таблиця 11 – Матриця відстаней.

Матриця відстаней										
	d1	d2	d3	d4	d5	d6	d7	d8	d9	d10
d1	0.0000	0.0117	0.0072	0.0101	0.0462	0.0542	0.0110	0.2168	0.2094	0.2272
d2	0.0117	0.0000	0.0045	0.0016	0.0579	0.0424	0.0007	0.2051	0.1977	0.2155
d3	0.0072	0.0045	0.0000	0.0029	0.0534	0.0470	0.0038	0.2096	0.2022	0.2200
d4	0.0101	0.0016	0.0029	0.0000	0.0563	0.0440	0.0009	0.2067	0.1993	0.2171
d5	0.0462	0.0579	0.0534	0.0563	0.0000	0.1003	0.0572	0.2630	0.2556	0.2734
d6	0.0542	0.0424	0.0470	0.0440	0.1003	0.0000	0.0431	0.1626	0.1552	0.1731
d7	0.0110	0.0007	0.0038	0.0009	0.0572	0.0431	0.0000	0.2058	0.1984	0.2162
d8	0.2168	0.2051	0.2096	0.2067	0.2630	0.1626	0.2058	0.0000	0.0074	0.0104
d9	0.2094	0.1977	0.2022	0.1993	0.2556	0.1552	0.1984	0.0074	0.0000	0.0178
d10	0.2272	0.2155	0.2200	0.2171	0.2734	0.1731	0.2162	0.0104	0.0178	0.0000

Критичне значення радіуса кулі:

r=
0,057624

Отже, перша підмножина містить такі об'єкти:

$A1=\{d1,d2,d3,d4\}$

$A2=\{d5\}$

$A3=\{d6,d7\}$

$A4=\{d8,d9,d10\}$

Проаналізуємо розподіл об'єктів в такому випадку. Для цього визначимо структуру ознак для кожного об'єкта, поділивши значення ознак. Результати для аналізу наведені в табл. 14.

Таблиця 14 – Структура даних.

Структура даних			
n	X	Y	X/Y
1	1.2	4.2	0.285714
2	1.3	4.4	0.295455
3	1.4	4.8	0.291667
4	1.5	5.1	0.294118
5	1.7	6.8	0.25
6	2.2	6.6	0.333333
7	2.3	7.8	0.294872
8	2.6	4.9	0.530612
9	2.7	5.2	0.519231
10	2.9	5.3	0.54717